

#eVALUate:

Monetizing Service Acquisition Trade-offs Using the

QUALITY-INFUSED Price[©] Methodology



*Capt Daniel J. Finkenstadt, USAF, and
Lt Col Timothy G. Hawkins, USAF (Ret.)*

The federal government persistently fails to make service contracts a managed outcome. Consequently, the three objectives of public procurement (transparency, value for money, and meeting requirements) are jeopardized. This research identifies the culprits as methodologies that are incompatible with the characteristics of services. These methodologies involve best-value source selection and contractor performance-information collection and evaluation. A new method of best-value proposal evaluation is offered that enables the buying agency to validly measure service quality, then to trade off levels of service quality with price, resulting in a Quality-Infused Price (QIP)[©]. The concept is tested on a task order competition using a case study methodology. Findings suggest that service quality can be monetized and that the application of a QIP[©] methodology can result in a superior sourcing decision. Additionally, fewer and higher quality proposals will be received. Based on the findings, conclusions are drawn and suggestions for future research are offered.



Over the past several decades, the United States transitioned from a goods-based to a services-based economy (McCullough, 2012; Powell & Snellman, 2004). As of 2013, services accounted for 78 percent of the country's gross domestic product and employed 82 percent of the country's workforce (U.S. International Trade Commission, 2015). At 68 percent of total contract spending in 2014, federal spending on services is substantial (Schwartz, Ginsberg, & Sargent, 2015). The Department of Defense (DoD) obligated 45 percent of its contract spend on services—an equal proportion as that spent on goods (Schwartz et al., 2015). While the DoD increasingly relies on defense contractors for services, it lacks the key elements at the strategic and tactical levels to make service contracts a managed outcome (U.S. Government Accountability Office [GAO], 2007a). Improving the tradecraft in services contracting has been a federal focus for some time (GAO, 2006; GAO 2007b; GAO, 2009b; Kendall, 2015; GAO, 2001a, 2001b).

Service has been defined as “the application of specialized competences (knowledge & skills) through deeds, processes, & performances for the benefits of another entity or the entity itself” (Vargo & Lusch, 2004, p. 2). The Federal Acquisition Regulation (FAR) defines a service contract as a “contract that directly engages the time and effort of a contractor whose primary purpose is to perform an identifiable task rather than to furnish an end item of supply” (subpart 37.101). Services are characterized as complex, heterogeneous, intangible (Apte, Ferrer, Lewis, & Rendon, 2006), perishable, and inseparable (Ellram, Tate, & Billington, 2007). First, the intangible nature of services renders specifications and customer expectations to be imprecise (Ellram et al., 2007). Second, services are, by nature, heterogeneous (Hawkins, Muir, & Hildebrandt, 2011). This is especially true of services with a high labor content, as performance will vary between providers and will likely differ between customers and with time required to deliver services (Parasuraman, Zeithaml, & Berry, 1985). Like providers, customers also lack a homogenous definition of service quality for many specified services (Hawkins, Berkowitz, Muir, & Gravier, 2015). Because of this, and since consistency in levels of performance from service personnel is difficult to attain, the level of quality that a service provider expects to deliver may vary greatly from the level of quality that the customer expects to receive (Parasuraman et al., 1985). Third, services are frequently perishable; unlike goods, services cannot be held or stocked in inventory. Whereas inventory policies for goods allow firms to buffer variability in future demand with safety stock, service providers must change service capacity to meet demand fluctuations (Ellram et al., 2007). The perishability of services



also presents challenges for inspection; service outcomes for many services can be inspected or evaluated only at the time of service performance (Hawkins et al, 2011). Given these perplexing challenges, how does the government validly leverage—not eliminate—the use of subjective service quality in both selecting contractors (i.e., reduce the risk of adverse selection) and motivating their performance (i.e., reduce the risk of moral hazard)?

To explore this question, this research supposes two axioms surrounding the objectives of procuring activities. First, the three primary objectives of public procurement are: transparency (Gilbert, Schapper, & Veiga-Malta, 2009), value for money (Gilbert et al., 2009), and meeting agency requirements. Second, the ability to procure services effectively and efficiently is desirable and in the public's best interest (Gilbert et al., 2009).

Agencies constantly trade these objectives based on risk that considers how best to meet agency requirements, gain value for money, and maintain a transparent process. These three objectives do not operate in a vacuum; rather, they are interoperable. If the agency can clearly and efficiently articulate requirements and evaluation methodologies, that clarity should increase transparency and, thereby, reduce the risk of bid protest (i.e., delayed meeting agency requirements). That clarity should also enable offerors to propose best in accordance with the agencies' needs (i.e., yield more value for money). Agencies must strategically assess the three objectives when they determine their source selection methodology along the best-value spectrum, ranging from lowest price technically acceptable (LPTA) to full trade-off of price and nonprice factors.

The ability to define, evaluate, select, award, and then manage service contracts is a problem that has garnered significant attention in the past decade, yet the many identified problems have not been resolved. Further, related problems in these areas and industry feedback have led to a call to “monetize” trade-offs to allow industry to understand the relative importance of evaluation factors in a manner that equates quality/performance to dollars of value (Kendall, 2015). The purpose of this research, therefore, is to provide a new approach that bridges the best-value continuum divide, optimizes the three primary objectives of public procurement, and delivers best value to the public sector by accounting for—rather than ignoring—the inherently subjective valuation of services.

Source Selection

FAR 15.304 identifies that price/cost, quality of the product or service, and past performance must be evaluated when contracting by negotiation. These criteria are considered across the best-value continuum spanning from LPTA to full trade-off in which the noncost factors may be significantly more important than cost/price. Agencies must know their requirements well enough to establish the best evaluation approach for source selection across this spectrum considering the three aforementioned objectives. LPTA generally compares to full trade-off as shown in Appendix A.

Federal acquisition and industry professionals have noted the following issues with LPTA versus full trade-off (Watson, 2015).

- Full trade-off evaluations may be too complex if workforce experience is low.
- Less procurement administrative lead time (PALT) is a driving factor in using LPTA.
- Evaluation criteria need to be better defined to industry. Industry needs to know the relative weight of cost/price to trade factors.
- Industry needs to know the level of performance to offer.
- Industry needs help understanding the competitive effects of a higher performance offer.
- There is a desire to avoid protests (Hawkins, Gravier, & Yoder, in press). Agencies must create meaningful evaluation discriminators.
- Industry can't determine the buying agencies' priorities.
- LPTA is perceived as "low cost/low quality."
- Cost risk does not equate to proposal risk.

The acquisition team's challenge is to find the optimal point within the best-value continuum to deal with these issues. The means by which federal agencies deal with the primary indication of quality and value for money—past performance—must first be explored.

Contractor Performance Rating System

The GAO asserted that contractor performance reports should be timely, accurate, and complete to allow federal procurement officials to make informed source selection decisions in the future. Despite persistent attention from the Office of Federal Procurement Policy (OFPP) and the GAO, Contractor Performance Assessment Reports (CPAR) continue to be plagued by a lack of reporting, untimeliness, incompleteness, and inaccuracies (GAO, 2009a, 2014b; Gordon, 2011). Agencies reported that workforce shortages, work priorities, time constraints, and difficulty obtaining timely feedback from other parts of the acquisition workforce are affecting reporting compliance (GAO, 2013, 2014b).

U.S. Federal Government agencies use the Contractor Performance Assessment Reports System (CPARS) to collect contractor performance information and use the Past Performance Information Retrieval System to access it.

U.S. Federal Government agencies use the Contractor Performance Assessment Reports System (CPARS) to collect contractor performance information and use the Past Performance Information Retrieval System to access it. Additional information that helps buyers reduce the risk of adverse selection is available from the Federal Awardee Performance and Integrity Information System. The CPARS scores contractors using a rating system of criteria including quality, schedule, cost control, management, and small business utilization (GAO, 2014b). To address the weaknesses of the CPARS, assessing “reputation attributes” (Blott, Boardman, Caday, Elliott, Griffin, Mastronardi, & Quinn, 2015) has been suggested to more closely align to commercial, “crowd-sourced” supplier performance evaluations. Such evaluations are updated in real time, known as “point-of-service,” as seen with online platforms such as Amazon, Yelp, Foursquare, etc. (Whetsell, 2015). Point-of-service platforms allow the customer to essentially “score” the vendors on their subjective experience, based on objective realities, in a timely manner that can lead to increased accuracy of reporting (Whetsell, 2015). Currently, CPARS reporting occurs annually with a 60-day contractor review window. Reports are not required until 120 days following the first 365-day period of performance. This means that it can take up to 485

days to officially capture the service quality delivered for an annual service (CPARS, 2015). Crowd-sourced, point-of-service reporting leads to a more holistic view of the contractor's performance in near real-time.

For instance, some customers are satisfied with a small proportion of late deliveries. Yet, others are upset with a contractor's inability to perform to all of the terms of the original contract. Both customers are receiving the same objective performance in the late delivery, but they may reach two different scores when rating the contractor. Many would view this level of subjectivity as a flaw in the rating system; however, the subjective reputation scoring embeds and assesses the contractor's ability to manage across all customers and demonstrates a truer measure of the contractor's customer management abilities. How well a contractor manages relationships across its market share of customers becomes apparent in this type of performance evaluation. This may best demonstrate the risk of partnering with the contractor on future service needs considering their ability to balance their customer relationship priorities. Such scoring methodologies account for the aforementioned characteristics of services—complexity, heterogeneity, and intangibility. This methodology may increase the chances of obtaining value for money in an efficient and effective manner while meeting the agency's requirements.

Quality-Infused Price (QIP)[©]

Tying the best-value source selection method and contractor past performance rating together "monetizes the trade-off" (Kendall, 2015). The source selection method used and the performance assessment method used must enable industry to understand the competitive effects of higher performance offers and to discern the level of performance to propose in response to a solicitation (Watson 2015). Benefits of monetizing the trade-off, aside from the expected better value offer, include faster PALTs and reduced protest risk. Recently, the Under Secretary of Defense for Acquisition, Technology, and Logistics, or USD(AT&L), pointed out that, although the number of DoD protests have increased in recent years (from 2001 to 2013), the sustainability rate of those protests has dramatically declined. The USD(AT&L) concluded that the Better Buying Power initiative to define value better in "best value" may be a significant contributor to that success (Defense Science Board Task Force, 2011). Monetizing trade-offs more clearly defines this value in terms of dollars. The proposed QIP[©] method aligns with and expounds upon the Better Buying Power initiatives.

The proposed QIP[®] concept addresses all three of the public procurement objectives—transparency, value for money, and meeting agency requirements. While the USD(AT&L)'s direction is heavily concerned with defining how much more buying agencies would pay for performance of a system or "thing," it falls short in defining monetized trade-offs for something as complex, heterogeneous, and intangible as a service. Evaluating services requires a midpoint between LPTA and full trade-off. Such a methodology should seek to give the combined benefit of faster evaluation processes (meeting agency requirements), more clear criteria to aid industry in deciding how to position the quality of their offer versus the costs of their offer (value for money), and an understanding of the agency's award decision (transparency).

Further, the QIP[®] methodology should not end with the award decision. Components of the methodology should be used with assessing contractor past performance in a way that becomes a program of record, or "score," for each firm. Consider an individual's credit score. Credit scores (e.g., FICO scores) quantitatively encapsulate past financial, contractual, and behavioral performance to indicate the risk of loss of lending to an individual. A similar model can be used by federal agencies to determine the quality risk, management risk, cost risk, and "other" risks related to trusting a particular firm—the firm's "reputation" currency.

To find the previously discussed midpoint on the best-value continuum, we propose the use of a composite Quality Adjustment Factor (cQAF) in developing a QIP[®] (i.e., an evaluated price adjusted for service quality) (Finkenstadt, 2015). Such a measure provides for faster PALT, more clear criteria for award, monetizes the trade-off for industry, and creates a past



performance standard that more closely aligns with the shift in commercial performance management. This system can open the door to new ways of conducting source selections while adding the post-award benefits found in incentive contracting to all forms of service procurement.

This system can open the door to new ways of conducting source selections while adding the postaward benefits found in incentive contracting to all forms of service procurement.

The cQAF described previously is a factor that may be greater to, equal to, or less than 1. It is derived from subjective service quality measures. The cQAF is used to assign a relative level of quality to the proposed price, considering factors determined to be germane to service value to the agency (Finkenstadt, 2015).

Once an offeror's prices are determined to be fair and reasonable, the agency applies the cQAF to the prices. Following the intent of FAR 15.304(c)(2), the agency would evaluate the quality of services being proposed by each offeror. To establish a value rating commensurate with the quality of the services being offered, the agency may use one or both parts of this two-part methodology, as shown in Table 1. The first part assesses relevant past performance, and consists of developing a composite Service Value Index (cSVI) using survey data from the offeror's previous customers (Finkenstadt, 2015). Contracts that leverage post-award incentives such as award fees and incentive fees can be considered in the establishment of this cSVI, either by having such ratings impact the score or subjectively as raters consider such factors in determining their level of satisfaction. This element remains to be codified and could depend on the type of service. The second part assesses the quality of the offeror's proposal considering relevant service quality indicators particular to the requirement such as personnel qualifications, technical process excellence, and management capability. The second part results in developing a composite Proposal Quality Rating (cPQR) (Finkenstadt, 2015).

TABLE 1. CQAF CATEGORY RATING TOOL

PRIMARY CATEGORY	SUBCATEGORY	Avg Survey Rating	Factor Score	Aggregate Score	Weight	Total Value Indices
Reliability	When the contractor's management promises to do something by a certain time, it does so	3	1.00	0.99	0.3	0.297
	When you have performance problems, the contractor's management is sympathetic and reassuring	3	1.00			
	The contractor is dependable	5	0.90			
	The contractor provides its services at the time it promises to do so	4	0.95			
Assurance	The contractor keeps accurate records	2	1.10	0.96	0.4	0.385
	This rater trusts the contractor's employees	3	1.00			
	This rater feels safe in its interactions with the contractor's employees	3	1.00			
	The contractor's employees are polite, professional, and courteous	4	0.95			
	The contractor's employees receive the support they need from the company to do their jobs well	5	0.90			

TABLE 1. CQAF CATEGORY RATING TOOL, CONTINUED

PRIMARY CATEGORY	RESPONSIVENESS TO REQUIREMENTS	Avg Survey Rating	Factor Score	Aggregate Score	Weight	Total Value Indices
The contractor met the requirements of the contract	The contractor satisfied our need(s)	4	0.95			
	The contractor performed the work we needed it to do	3	1.00			
	The contractor's work was timely	3	1.00			
	This rater was satisfied with the quality of the contractor's work	3	1.00			
SUBCATEGORY	COMPOSITE SERVICE VALUE INDEX (CSV) ADJUSTMENT FACTOR				1	0.976
PRIMARY CATEGORY	PROPOSAL PROCESS EFFICIENCY AND RESOURCE ALLOCATION PLAN				1.01	0.5
How satisfied are you with the quality design of process flows proposed?	How satisfied are you with the quality of personnel capacity management proposed?	3	1.00			
	How satisfied are you with the quality of demand planning processes proposed?	3	1.00			
	How satisfied are you with the quality of continuous process improvement proposed?	2	1.10			

TABLE 1. CQAF CATEGORY RATING TOOL, CONTINUED

PRIMARY CATEGORY	Staffing Quality	Avg Survey Rating	Factor Score	Aggregate Score	Weight	Total Value Indices
SUBCATEGORY	How satisfied are you with the quality of staffing proposed for each line item?	5	0.90	0.90	0.5	0.45
	Composite Proposal Quality Rating (CPQR) Adjustment Factor			1	0.95625	
cQAF	Final Composite Rating	Weight	Total Adjustment Factor			
CSV	0.976	0.3	0.2928			
CPQR	0.95625	0.7	0.669375			
cQuality Adjustment Factor for QIP[©] Calculation			0.962175			

Note. Notional, including weighted importance

To establish the final cQAF, cSVI, and cPQR, first a scaled rating system that converts subjective service quality into objective factors is needed. This is the moment in which the trade-off is monetized. As such, these factors should not be established arbitrarily. Each type of service should be investigated using market research to determine the appropriate amount of “value for money” that each level of service quality represents to a majority of customers. This value-for-money scale may be created through market research into leading performance indicators in a particular type of service. A simplified five-point Likert scale is offered in Table 2. Note that the scaling creates “golf-like” reverse indices that increase with negative ratings and decrease with positive ratings (Finkenstadt, 2015).

TABLE 2. CQAF RATING SCALE (NOTIONAL)

Adjective Rating	Numerical Rating	cQAF
Strongly Disagree	1	1.1
Somewhat Disagree	2	1.05
Neither Agree Nor Disagree	3	1.0
Somewhat Agree	4	0.95
Strongly Agree	5	0.90
Not Rated	None	Not Included

The cSVI is the factor that would become the crowd-sourced reputation score (i.e., the “numerical rating” listed in Table 2). A cSVI survey should be developed using an established scale with valid psychometric properties. The service quality scale included in Table 2 was recently developed for a business-to-business context (Hawkins et al., 2015), but may need further refinement by type of service (i.e., design-engineering services, testing services, facility management services, etc.). These assessments would be subjective in nature and are intended to systematically capture the quality of a particular firm operating within the type of service as assessed by the most recent and relevant customers. This assessment would be solely at the agency’s discretion in determining best value for each requirement. This part of the cQAF could replace the fallible (Blott et al., 2015) CPARS. The cSVI could be used for near real-time ratings that, even if constrained by the current vendor 60-day review window, would reduce final service performance reporting by up to 88 percent when compared to the maximum CPARS annual reporting window of 485 days (CPARS, 2015).

The cPQR is unique to each acquisition and may or may not be used in addition to the cSVI to establish the cQAF. It should be established using questions for the technical/quality evaluation team members to consider in scoring each proposal. This would be similar to the areas that are considered significant technical subfactors within a proposal. The agency would then derive the final cQAF to be used to establish the QIP® by combining the cSVI and cPQR factors using an agency-determined weight of importance per factor. These factors can be combined to yield a single cQAF for adjustment or may be used independently as the sole QIP® adjustment factor (Finkenstadt, 2015). This process may become agency- and/or service industry-dependent, and should be considered by agencies prior to implementation.

Once the agency calculates the cQAF for each offeror, the agency would apply the cQAF to the total price of each line item within the offeror's proposal (Finkenstadt, 2015). The agency would then award to the conforming offeror demonstrating the best-quality offer in terms of both price and quality ratings—in other words, the lowest evaluated QIP® offer.

An example involving advisory services is shown in Table 3. In this example, the cQAF of 0.962 is derived from the calculations in Table 1 by rolling up a notional cSVI at a relative importance weighting of 30 percent and a notional cPQR at a relative importance weighting of 70 percent. The 30 percent weight on cSVI and 70 percent weight on cPQR are notional; the agency would determine these weights depending on what is more important—actual service quality from past work or promises of future service quality in the proposal. The cSVI rating of 0.976 is created by weighting the scores of each primary factor of reliability, assurance, and responsiveness to requirements. The cPQR rating of 0.956 is derived by weighting the firm's process management plan and staffing quality. Again, all weightings are notional and would be established prior to developing the final cQAF. Using this example, the offer would be assessed as having an inherent quality value of \$161,647.79. Award would be made for \$4,273,570.00, but the offeror would get "credit" for having a lower proposed price based on carrying a higher quality rating (cQAF). This is the final step in "monetizing" the trade-off (Finkenstadt, 2015).



TABLE 3. OFFER QUALITY-INFUSED PRICE (QIP)® CALCULATION TABLE (NOTIONAL)

QIP® Calculation Example

	Quantity	Units	Unit Firm Fixed Price	Total Firm Fixed Price
CLIN 0001AA (Base)	25	Lot	\$4,000.00	\$100,000.00
CLIN 0002AA (Base)	25	Lot	\$24,000.00	\$600,000.00
CLIN 0001AB (Option 1)	25	Lot	\$4,400.00	\$110,000.00
CLIN 0002AB (Option 1)	25	Lot	\$26,400.00	\$660,000.00
CLIN 0001AC (Option 2)	25	Lot	\$4,840.00	\$121,000.00
CLIN 0002AC (Option 2)	25	Lot	\$29,040.00	\$726,000.00
CLIN 0001AD (Option 3)	25	Lot	\$5,324.00	\$133,100.00
CLIN 0002AD (Option 3)	25	Lot	\$31,944.00	\$798,600.00
CLIN 0001AE (Option 4)	25	Lot	\$5,856.40	\$146,410.00
CLIN 0002AE (Option 4)	25	Lot	\$35,138.40	\$878,460.00
Total Proposed Price			\$4,273,570.00	
Total Proposed Price CLIN 0001 & Options				\$610,510.00
Total Proposed Price CLIN 0002 & Options				\$3,663,060.00
cQAF			0.962175	
Quality-Infused Price (QIP)® CLIN 0001 & Options				\$587,417.45925
Quality-Infused Price (QIP)® CLIN 0002 & Options				\$3,524,504.75550
QIP® All CLINs & Options				\$4,111,922.21
TOTAL ASSESSED VALUE				\$161,647.79

Case Study

A case study methodology was used to test a portion of the cQAF in a recent source selection for administrative support services. This task order included two line items for each year of a 5-year service contract: one line item for contract support services and a second, larger line item for program control (financial) analysis services. The case involved evaluating offerors for task order awards under a prepriced indefinite delivery–indefinite quantity (IDIQ) contract in which 11 offerors could offer better than on-contract pricing, but had to offer no higher than on-contract pricing. The request for proposal allowed the agency to decide on awarding task orders for one or all of the line items. The IDIQs did not allow for past performance evaluations in the base year of the IDIQs, as the agency considered past performance during the base award to be at least satisfactory for all contractors and prohibited further past performance evaluations until the end of the IDIQ base period. This meant that all trade-offs for nonprice factors could not utilize past performance; therefore, the cSVI could not yet be tested. The agency chose this IDIQ for a limited test case due to (a) perceived weaknesses in proposal quality, (b) post-award performance results on recently LPTA-awarded task orders, and (c) a low threat to mission if the evaluation methodology were found to be flawed or was contested.

Only three of the 11 IDIQ holders submitted a proposal. The overall assessed quality of these three proposals was relatively high compared to historical LPTA evaluations for similar services. The lowest priced offeror was not selected due to having the lowest cPQR quality rating. Since the agency stated that quality was considered more important than price, and it did not intend to enter discussions, award was made to the highest rated offeror in terms of quality. All offerors proposed pricing at or below those listed in the base IDIQ. The lowest offeror's pricing was perceived as questionably low and would have driven the team into discussions had their quality rating been higher. This source selection did not fully apply the concept of adjusting evaluated pricing since it was a first trial. The researcher first wanted to determine whether the quality rating system would affect the quality trade-off.

In this case, the highest priced yet highest rated offeror was selected, while the overall price remained 4 percent below the agency's estimate. The cPQR method allowed for a team of three personnel to assess three full proposals in only 3 days. Quality perceived was converted to a rating that yielded the results shown in Table 4. Actual cPQR scores are not available due to source selection material and the sensitive nature of the procurement. However,

this table demonstrates their relative placement after applying the cPQR as a general quality ranking independent of QIP[®] adjustment. The full QIP[®] methodology, applying the cQAF to price in order to rank offerors based on QIP[®], was not completed in the actual source selection. The agency surmised that limited application of the methodology would minimize industry confusion over an unfamiliar evaluation methodology, but would allow for early testing of the concept. This case was a first-off trial, and this research calls for agencies to consider future and full application of the QIP[®] methodology.

TABLE 4. CASE STUDY TRADE-OFF EVALUATION

Rating Factor	Offeror X	Offeror Y	Offeror Z*
Price (CLIN 0001)	1 of 3	3 of 3	2 of 3
Price Variance from Low	0%	37%	28%
Price (CLIN 0002)	1 of 3	2 of 3	3 of 3
Price Variance from Low	0%	14%	25%
Nonprice (cPQR) CLIN 0001	3 of 3	2 of 3	1 of 3
Nonprice (cPQR) CLIN 0002	2 of 3	3 of 3	1 of 3

*Awarded offeror

Results

Although no evaluated price calculations were made during the actual source selection using QIP[®], the scaling methodology was maintained to allow the researcher to take actual source selection data and run the scenario utilizing the full QIP[®] scoring concept to identify strengths and weaknesses in the concept post hoc so as not to affect the actual award decision. In this application, the researcher applied a cPQR created by using the program management plan and the staffing quality criteria similar to Table 1, yet heavily customized for each type of service being procured. The weightings of the plan versus staffing quality are source selection-sensitive, but offerors were told which held the greatest importance to aid in proposal development.

Quality trade-offs in a trial service contract source selection resulted in useful measures of service quality, cooperative industry participation, fewer and higher quality proposals (i.e., less risk of adverse selection and greater efficiency), and a different contractor selection than a typical subjective price-performance trade-off. The effectiveness of this QIP[®] methodology

must be supported by scaling that considers relative price variations in the market. Had QIP® been applied to the actual source selection, the lowest quality offeror having the lowest evaluated price may have been selected because the quality scaling was not sufficient to overcome the wide variation in pricing. Discussions would have been necessary to determine the reasonableness of the lowest offeror's pricing to reduce the risk of "buying-in." These discussions would have been completed prior to applying the QIP® to ensure that the final results were accurate.

When applying QIP®, the test case shows that the selected offeror becomes more competitive in terms of evaluated price, relative to the low, based on evaluated quality, and does in fact, displace the second lowest offeror from the non-QIP® evaluation (Table 5). Thus, the QIP® methodology demonstrates the ability to drive value for money into an evaluation and to create source selection results that more closely align to traditional, yet more subjective, full trade-off methods in a rapid manner that is more transparent and easier to use.

TABLE 5. CASE STUDY QIP®-ADJUSTED TRADE-OFF EVALUATION

Rating Factor (After QIP®)	Offeror X**	Offeror Y	Offeror Z*
QIP® (CLIN 0001)	1 of 3	3 of 3	2 of 3
Price Variance from Low	0%	33% ↓	16% ↓
QIP® (CLIN 0002)	1 of 3	3 of 3 ↓	2 of 3 ↑
Price Variance from Low	0%	17% ↑	12% ↓

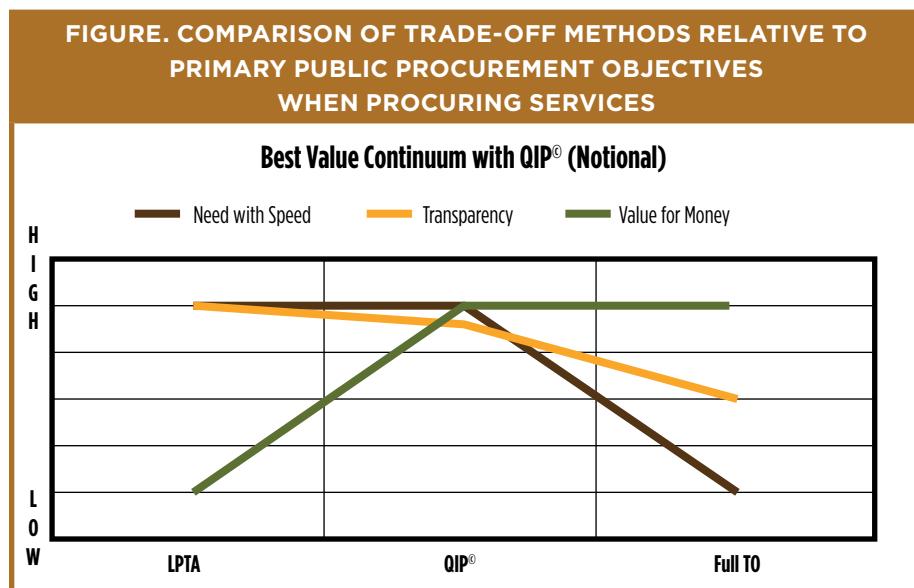
*Awarded offeror

**QIP® best-value offeror

To validate results, a questionnaire was sent to all quality team evaluators postaward (Appendix B). The responses indicated that all evaluators found the cPQR methodology easy to use, easy to understand, asked the right types of questions, and resulted in the best value to the government. The only area listed for cPQR improvement related to requiring the evaluation team to have earlier and more robust input into the relative importance weighting of cPQR categories. The ease of use and ability for lesser trained personnel to administer this methodology show significant promise in reducing the risk to poorly executed, best-value trade-off evaluations that can occur due to less experienced evaluators (Watson, 2015).

Discussion

The Figure demonstrates that while LPTA provides for a faster PALT and is relatively transparent, it sacrifices value for money in service acquisitions. While full trade-off has the capability to maximize value for money, it may reduce transparency if evaluations become too complex, and most assuredly will sacrifice speed of the service acquisition. QIP[®], as proposed, would provide a means for monetizing trade-offs. Monetizing trade-offs prevents pre-award questions related to full trade-off ambiguity as well as post-award delays due to protest. The QIP[®] provides for faster acquisition of needs, with a clearer evaluation methodology and trade-offs that increase both value for money and transparency.



The use of QIP[®] has the potential to improve or eliminate major gaps found in the current best-value source selection process (Watson, 2015) for services. QIP[®] provides the following:

- Monetized trade-offs (Transparency)
- Ability to pay more for service quality when prudent (Value for Money)
- Clear communication of federal agency priorities in price and nonprice factors to offerors (Transparency)

- May help to correct for wide price disparities in previously negotiated multiple-award contracts (i.e., when the low would otherwise always win if LPTA were the only option to full trade-off) (Value for Money)
- Rapid evaluation and acquisition capability (Meeting Requirements/Need with Speed)
- Clear evaluation criteria that reduce protest risk (Transparency)

In addition, the QIP® cSVI component, as a crowd-sourced form of past performance, has several advantages:

- Encourages higher compliance rates for past performance reporting by providing a clear, easy-to-use format with more resemblance of commercial, crowd-sourced contractor performance reporting
- Fills past-performance assessment repository gaps
- Promotes rapid evaluation and acquisition capability (could replace the entire past performance volume requirement in proposals)
- Encourages better life-cycle performance with contractors (i.e., contractors with lower cSVIs will have price advantages and can offer higher quality services assuming a better QIP®, while higher cSVI contractors will have to be more aggressive in pricing in the near-term and improve quality in the long-term to keep market share and realize higher future returns)
- Encourages pricing off-sets for performance issues
- Creates clear discriminators for services based on customer ratings (subjective customer quality is a truer way to assess the intangibility inherent in service performance)
- Arms federal agencies with real-time market performance data
- Enables more accurate and more efficient supplier ranking (e.g., DoD superior supplier incentive program). More efficient rankings will enable rankings by type of service rather than be limited to the top 30 business units by dollars obligated annually (DoD, 2015).

Conclusions

A QIP[®] methodology using an established cSVI system shows great promise in progressing the state of the art in contractor performance management while finding a desirable midpoint along the best-value continuum. This research calls for federal agencies to consider adopting such a methodology to meet public procurement objectives. The QIP[®] and cSVI concepts may be seen as “lofty” or even naïve from a historical federal procurement policy vantage point. However, that vantage point is built upon a history of ill-fated service contractor rating systems that never meet the intent of federal agencies to improve transparency, value for money, and requirements satisfaction in highly efficient and effective means. The current CPARS has more focus on getting the reports completed versus the accuracy and value of the reports—particularly for its intended purpose of better informing future source selections. Understanding the higher level impacts of the system as it relates to transparency, value for money, and meeting agency requirements should be the ultimate goal of any contractor performance rating system as well as the source selection process it feeds.

Understanding the higher level impacts of the system as it relates to transparency, value for money, and meeting agency requirements should be the ultimate goal of any contractor performance rating system as well as the source selection process it feeds.

Critics may question the ability to adjust an offeror’s evaluated price based on subjective evaluation inputs. However, the government does this today with the concept of most probable price and cost evaluations to determine what the agency anticipates the actual cost or price of an offer will be considering all risks. Considering that service quality is a primary risk concern in a services acquisition, the concept of QIP[®] is not a radical idea. Others question the idea of crowd-sourcing something as sensitive as contractor performance for federal contract award decisions. This can be mitigated by controlling the “crowd” as we do today with CPARS. Agencies should ensure that only contracting officers, contracting officer technical representatives, and possibly program managers have access to the cSVI rating system.

Limitations and Future Research Directions

This study is not without limitations. First, it is a limited application of one case. Future research could expand the number and variety of cases of application. Future research employing a quasi-experiment could compare sourcing and performance (i.e., the full service life cycle) of multiple service procurements of the same type of service to examine differences in value and service quality. Further research should also explore the customization of dimensions of the business-to-business service quality measurement scale. Different types of services will likely be more validly measured by customized aspects of service quality. Additionally, since different services span a vast spectrum of scope and complexity, further research could explore whether the proposed QIP® methodology will be equally effective across the different types of services. In closing, this article serves as a call to agencies to pilot-test the QIP® concept.



References

Apte, U., Ferrer, G., Lewis, I., & Rendon, R. G. (2006). *Managing the service supply chain in the Department of Defense: Opportunities and challenges* (Report No. NPS-AM-06-032). Monterey, CA: Naval Postgraduate School.

Blott, J., Boardman, N., Cady, A., Elliott, J., Griffin, W., Mastronardi, N., & Quinn, P. (2015). *Auction and game theory based recommendations for DOD acquisitions* (Report No. USAFA-AM-14-189). Acquisition Research Program Sponsored Report Series. Monterey, CA: Naval Postgraduate School.

Contractor Performance Assessment Reporting System. (2015). *Guidance for the Contractor Performance Assessment Reporting System (CPARS)* (December 2015). Retrieved from: <https://www.cpars.gov/pdfs/CPARS-Guidance.pdf> on March 23, 2016.

Defense Science Board Task Force. (2011). *Report of the Defense Science Board Task Force on improvements to services contracting*. Washington, DC: Office of the Under Secretary of Defense (Acquisition, Technology, and Logistics).

Department of Defense. (2015). *Performance of the Defense Acquisition System, 2015 annual report*. Washington, DC: Office of the Under Secretary of Defense (Acquisition, Technology, and Logistics).

Ellram, L. M., Tate, W. L., & Billington, C. (2007). Services supply management: The next frontier for improved organizational performance. *California Management Review*, 49(4), 44–66.

Finkenstadt, D. J. (2015, June 9). *NRO quality evaluations & “total evaluated price”* [Powerpoint slides]. Presentation to director, Office of Contracts, National Reconnaissance Office, Chantilly, VA.

Gilbert, D., Schapper, P. R., & Veiga-Malta, J. N. (2009). Framework for the management and reform of public procurement. In K. V. Thai, *International handbook of public procurement* (Chap. 4), CRC Press, Boca Raton, FL.

Gordon, D. I. (2011). *Improving contractor past performance assessments: Summary of the Office of Federal Procurement Policy’s review and strategies for improvement* [Memorandum]. Washington, DC: Office of Federal Procurement Policy, Office of Management and Budget.

Hawkins, T., Berkowitz, D., Muir, W., & Gravier, M. (2015). Improving services supply management in the defense sector: How the procurement process affects B2B service quality. *Journal of Purchasing and Supply Management*, 21(2), 81–94.

Hawkins, T., Gravier, M., & Yoder, E. C. (in press). Federal bid protests: Is the tail wagging the dog? *Journal of Public Procurement*, 16(1).

Hawkins, T., Muir, W., & Hildebrandt, G. (2011). Determinants of services sourcing performance. *Proceedings of the 8th Annual Acquisition Research Symposium* (Vol. 1, Report No. NPS-AM-11-017, pp. 212-234), Naval Postgraduate School, Monterey, CA, May 11-12.

Kendall, F. (2015). *Implementation directive for Better Buying Power 3.0 – Achieving dominant capabilities through technical excellence and innovation* [Memorandum]. Washington, DC: Office of the Under Secretary of Defense (Acquisition, Technology, and Logistics).

King, R. (2014). *Biometrics and national security white paper*. Biometrics Research Group, Inc. Retrieved from <http://www.academia.edu/7434174/Biometrics>

McCullough, N. (2012). *Services, not manufacturing, will revive the U.S. workforce*. Retrieved from <http://www.forbes.com/sites/ciocentral/2012/04/19/services-not-manufacturing-will-revive-the-u-s-workforce/#6ec7e6da6443>

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49(1), 41-50.

Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology*, 30(1), 199-220.

Schwartz, M., Ginsberg, W., & Sargent Jr., J. F. (2015). *Defense acquisitions: How and where DOD spends its contracting dollars* (Congressional Report No. 7-5700). Washington, DC: Congressional Research Service.

U.S. General Accounting Office. (2001a). *Contract management: Improving services acquisitions* (Report No. GAO-02-179T). Washington, DC: Author.

U.S. General Accounting Office. (2001b). *Contract management: Trends and challenges in acquiring services* (Report No. GAO-01-753T). Washington, DC: Author.

U.S. Government Accountability Office. (2006). *Defense acquisitions: Tailored approach needed to improve service acquisition outcomes* (Report No. GAO-07-20). Washington, DC: Author.

U.S. Government Accountability Office. (2007a). *Defense acquisitions: DoD needs to exert management and oversight to better control acquisition of services* (Report No. GAO-07-359T). Washington, DC: Author.

U.S. Government Accountability Office. (2007b). *Defense acquisitions: Improved management and oversight needed to better control DOD's acquisition of services* (Report No. GAO-07-832T). Washington, DC: Author.

U.S. Government Accountability Office. (2009a). *Federal contractors: Better performance information needed to support agency contract award decisions* (Report No. GAO-09-374). Washington, DC: Author.

U.S. Government Accountability Office. (2009b). *Defense acquisitions: Actions needed to ensure value for service contracts* (Report No. GAO-09-643T). Washington, DC: Author.

U.S. Government Accountability Office. (2013). *Contractor performance: DOD actions to improve the reporting of past performance information* (Report No. GAO-13-589). Washington, DC: Author.

U.S. Government Accountability Office. (2014a). *Defense contracting: Factors DoD considers when choosing best value processes are consistent with guidance for selected acquisitions* (Report No. GAO-14-584). Washington, DC: Author.

U.S. Government Accountability Office. (2014b). *Contractor performance: Actions taken to improve reporting of past performance* (Report No. GAO-14-707). Washington, DC: Author.

U.S. International Trade Commission. (2015). *Recent trends in U.S. services trade: 2015 annual report* (Publication No. 4526). Washington, DC: Author.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, 68(1), 1-17.

Watson, K. (2015). *LPTA versus tradeoff: How procurement methods can impact contract performance*. Naval Postgraduate School Joint Applied Project. Monterey, CA: Naval Postgraduate School.

Whetsell, J. (2015, August). Vendor performance management: Improving the collection and use of performance information. *Contract Management*, 41-47.

Appendix A

Comparison of Best-Value Source Selection Options Relative to Primary Public Procurement Objectives

Objective	LPTA	Full Trade-off	Rating Rationale
Transparency	High	Low	<ul style="list-style-type: none"> • LPTA typically defines evaluation criteria in a very clear and objective fashion. Protest risk is minimized if LPTA process is followed. • Full trade-off may allow for high levels of subjectivity at the factor and subfactor level, and runs the risk of being challenged both pre- and post-award.
Value for Money	Low	High	<ul style="list-style-type: none"> • LPTA clearly states the agencies' desire to pay less for a base requirement and no more. May drive "bare minimum" solutions from industry in an effort to remain competitive. High risk of "buying-in." • Full trade-off establishes areas of trade that are primarily quality- and performance-based; reduces the risk of post-award performance issues (GAO, 2014a).
Meeting Agency Requirements (Need with Speed)	Low	Medium	<ul style="list-style-type: none"> • LPTA tends to meet timelines and basic requirements. Can be risky if unknowns surface postsolicitation. • Full trade-off expends the greatest amount of time in order to minimize the risk of unknowns insofar as the evaluation criteria plans for it (i.e., assessing proposal risk). However, minimizing unknowns equates to unclear subjectivity that may increase protest risks.

Note. High = highest level relative to alternative; Med = essentially the same as the alternative; Low = lowest level relative to alternative.

Appendix B

Post-award cPQR Evaluator Questionnaire

1. How easy/hard did you find the criteria to understand?
2. Were the evaluation tools easy to use or hard to use?
3. Did we [the agency] ask the right questions in the evaluation or could we have done better?
4. Is there anything you would do to improve this evaluation method in the future?

Appendix C

Post-award cPQR Vendor Questionnaire

1. Was your decision to propose or “no bid” [actual term used within the ordering procedures of the base IDIQ] based on the cPQR methodology used?
2. What, if anything, did you change about your traditional proposal methods in order to meet the requirements of this request for proposal’s cPQR methodology?
3. Was the cPQR evaluation methodology easy to understand?
4. What, if anything, would you change about the cPQR evaluation methodology used?
5. Considering your experience, in the future would you be open to having your evaluated price* adjusted based on the score received using a similar evaluation methodology?
6. Did the cPQR methodology encourage your firm to focus more on price or nonprice (i.e., quality) factors in proposing?

*Note. The price would be adjusted for evaluation purposes only. The final award price would be as proposed or negotiated with the [agency].

Appendix D

Findings and Lessons Learned

Finding	Lessons Learned
The cPQR primary factors and subfactor areas and questions were highly indicative of a quality proposal from the requiring agencies' point of view	Procuring agencies should develop cPQR satisfaction questions in tight coordination with the requiring activity. Standardized, valid cPQR measures customized to each type of service could evolve over time.
The scaling factor used ranged from 0.85 to 1.15, but was not indicative of the potential pricing variations across the service line-item disciplines being proposed, and was based more on what the procuring agency calculated were rational price variations in typical procurements.	Procuring agencies should develop the scaling factor ranges based on market research into the commercial market's typical price variation across each service type, and not assume what a "fair" scaling should be. Note: This confirmed the risk to utilizing QIP® price adjustments when the scaling has not been developed based on robust market data.
Industry did not question the unique quality evaluation methodology.	Draft request for proposal (RFP) documents were posted to the business opportunity Web site to gather questions from industry and ensure it understood the methodology prior to issuing a final RFP. Draft RFPs and industry engagement are key when introducing new evaluation methods.
Only roughly a third of contractors on the Multiple Award IDIQ proposed to provide these services to the agency. Most no-bid letters received by the agency stated an inability to source personnel who met the quality requirements of the RFP. Anecdotal comparisons of number of offers received on similar RFPs under this IDIQ showed that LPTA yielded higher response rates with lower quality offers.	The agency received offers from only those contractors who could meet the agencies' desired quality needs. The natural desire found in LPTA to "buy-in" was minimized. The risk of adverse selection was mitigated and the selection was more efficient.

Appendix D, Continued

Finding	Lessons Learned
<p>The agency would have reached a different award decision had the QIP® been utilized versus leaving the final trade-off to a subjective comparison of cPQR scores versus prices offered.</p>	<p>Agencies must ensure they have robust data to support the cQAF primary categories and subcategory questions, and a solid understanding of the scaling of each rated area and the associated weights to create meaningful discrimination between offers within a service type. Further, the use of cSVI would have had additional effects on the final scoring and should be considered in all future QIP® source selections where past performance is being evaluated.</p>

Biographies



Capt Daniel J. Finkenstadt, USAF, is a procuring contracting officer serving at the National Reconnaissance Office (NRO) in Chantilly, Virginia. He has 12 years' contracting experience and is responsible for one of NRO's largest major systems acquisitions. Capt Finkenstadt has authored six publications related to contract management. He holds a BS in Business Management from Mt. Olive College and an MBA in Strategic Sourcing from the Naval Postgraduate School.

(E-mail address: asuscrash@gmail.com)



Lt Col Timothy G. Hawkins, USAF (Ret.), is an assistant professor at Western Kentucky University. He has published 18 articles exploring procurement and logistics in scholarly journals. Lt Col Hawkins holds a PhD in Marketing from the University of North Texas, an MS from Wilmington College, and a BS in Civil Engineering Technology from Western Kentucky University. While on active duty, he served as a warranted contracting officer for the U.S. Air Force.

(E-mail address: timothy.hawkins@wku.edu)